A Framework for Writing Measurement Requirements and its Application to the Planned Europa Mission

Sara Susca
Jet Propulsion Laboratory,
California Institute of Technology
4800 Oak Grove Dr.
Pasadena, CA 91109
Sara.Susca@jpl.nasa.gov

Laura L. Jones-Wilson
Jet Propulsion Laboratory,
California Institute of Technology
4800 Oak Grove Dr.
Pasadena, CA 91109
Laura.L.Jones@jpl.nasa.gov

Bogdan V. Oaida
Jet Propulsion Laboratory,
California Institute of Technology
4800 Oak Grove Dr.
Pasadena, CA 91109
Bogdan@jpl.nasa.gov

Abstract— **Science-engineering communication is critical to the success of any science-driven mission. The process of building this understanding relies on a shared language for communicating science needs and engineering results, which can be particularly difficult on large space-science missions where many different institutions contribute to the science team. The Science Traceability Matrix can be used to formalize this communication pathway, but it has limited use in the development of the science requirements flow down, and vary in format, scope and content from mission to mission. There are many guidelines on developing well-constructed requirements in general, but very little is published on how to actually write these science-driven requirements in a systematic way. This paper discusses the measurement-domain science traceability and alignment framework, or M-STAF, which was developed to help frame the conversation between scientists and engineers in the development of science measurement requirements. The M-STAF provides a common language that can be used to ensure consistency across instruments, completeness in the coverage of the requirements, and traceability of the engineering work to the science objectives of the project. This work discusses the framework in the context of other communication tools, how it can be implemented on a flight project, and provides examples of how it might be used to improve the measurement requirements set for a project. The general framework is presented through the lens of its potential application on the planned Europa Mission.**

TABLE OF CONTENTS

1. MOTIVATION AND INTRODUCTION

Requirement Types and Levels

For large science-driven space missions, one of the most important tasks of the systems engineering process in Phase A is to successfully translate the science mission objectives into an implementable set of engineering requirements. Indeed, it is expected that project systems engineers decompose the level 1 (L1) requirements (negotiated with the external customer) into lower level requirements which, over multiple levels, generate instrument performance requirements, mission requirements, spacecraft requirements, etc. These levels are defined by the project systems engineering team in order to capture a number of different types of information (responsibility, scope, traceability, etc.) but very little formal guidance appears in published literature.

Although every project develops their own assessment of which requirement types are categorized into which project level, missions (including the planned Europa Mission) can decompose L1 requirements into a set of *science requirements* which necessarily live at the level just below the customer requirements. For clarity, we will call these level 2 (L2) requirements. Their purpose is to fully constrain the scope of the science in a more complete and detailed way than may be captured in the requirements from the external customer. These requirements are written by members of the project science team and need the buy-in of the project science office and the project system engineering office before they can be used as parents to requirements at lower levels.

From here, there are many ways to reach the functional requirements for the instruments, the mission design, the spacecraft, etc. One possible route is to decompose the science requirements into *measurement requirements*. The measurement requirements need to quantitatively specify which observations must be collected to support a given science requirement. Hierarchically, these requirements exist as children to the L2 science requirements, however they can either live at many different levels, depending on the preference of the system engineering team. Regardless, the measurement requirements must communicate the constraints and scope on the observations that must be collected to support the science requirements. These requirements might include information on the conditions under which the observations must be taken (i.e., define what makes a valid measurement), the frequency and spatial distance the observations must be collected, the quality of the measurements, and any interconnections between an observation and the full complement of data.

978-1-5090-1613-6/17/$31.00 ©2017 IEEE

Challenges in Developing the Measurement Requirements

There are many reasons why generating the measurement requirements is so challenging. For one, writing these requirements takes a significant amount of work because no one stakeholder has all of the information necessary to compose well-formed and complete set. Although it is typically an instrument science representative that authors these requirements, there are many engineering stakeholders across the project that need to trace their design to specific measurement characteristics.

Secondly, these stakeholders often do not have a clear way of communicating their needs to the requirement authors. Consequently, the measurement requirements are usually quite clear to the investigation or instrument team that they serve, but are less so to the project engineers. To give a better idea of the diversity of the stakeholder set, consider that the mission design, flight system, payload, project system engineering teams, and, of course, the instrument scientists or principal investigators (PIs) that generated them must all agree on the meaning of these requirements. Each of these aspects of a project have different concerns that drive their technical and budgetary decision-making processes, but these sometimes-competing priorities are further exacerbated by frequent misunderstandings caused by the lack of a common vocabulary. The different elements of the team have overlapping technical vocabularies that – when used in requirements – lend themselves to overloaded terms and misinterpretations because each community uses the same language differently. For example, asking each community to define a "dataset" or a "campaign" leads one to marvel at the creative variety of the (all perfectly legitimate) uses of the term. The use of a common framework for communicating across these boundaries can help clear this confusion as long as all of the communities understand the language being used.

Finally, the sheer number of stakeholders involved – each using their own language to communicate their needs – means that generating a consistent and complete set of measurements requirements is especially difficult. The language confusion is compounded when the scope of the requirements varies across instruments. Often engineers find that requirements they need to fully specify their design are missing or unevenly addressed in the requirement flow-down, and the process of checking to ensure that each instrument has been fully specified at the same level is time-consuming and instrument-specific. Worse, not performing this task may mean that some measurement needs and therefore some science needs are not properly communicated across the project. A well-structured requirement generation process can address many of these concerns, but there are few available tools or recommendations on best practices for solving these problems early enough in the project lifecycle to be effective.

Science Traceability and Alignment Framework

The authors sought to address these issues by developing a novel approach for organizing the information conveyed in the science and measurement requirements called the Science Traceability and Alignment Framework (STAF). Much like the musical staff, the framework acts as an organizing set of principles that coordinates the conversation among many different stakeholders. To do this, the STAF sets up a common language and an associated set of tools that can be adopted in order to improve traceability, consistency, and completeness across these requirements sets. The project-level implementation of the framework, or P-STAF addresses the development of the science requirements while the measurement-level implementation, or M-STAF, addresses the measurement requirements development. Although both levels of the STAF work together, each level has different end-users and might require negotiations involving different people (and levels) within the project. Thus, the split between the M-STAF and P-STAF is preserved in this discussion of the architecture to ensure that the focus of each level is appropriate to those different users. This paper focuses on the M-STAF, whereas the companion paper [1] focuses on the P-STAF.

The remainder of this paper explains the current gaps in the available tools and then describes the STAF taxonomy with an emphasis on M-STAF-specific fields. We then address the uses and the implementation of the M-STAF. We also discuss the implications of this approach to model-based systems engineering approaches and verification and validation efforts. Throughout, we use the planned Europa Mission as a concrete example that underpins the abstract concepts being discussed.

2. State of Practice

IEEE standards define the characteristics of a good/valid set of requirements as complete, consistent, affordable, and bounded. [2]. Similarly, a valid single requirement must be unambiguous, consistent, complete, traceable, and necessary. [2] As noted above, these qualities can become difficult to achieve (and enforce) in the context of a measurement requirement flow down because of these competing stakeholders and communication issues. So what tools *are* available to the systems engineer who is trying to work with the science office to create a valid flow down of the measurement requirements? There are at least three different potential sources of guidance: 1) published standards and best practices, 2) the flow down of requirements on previous missions, and 3) the Science Traceability Matrix (STM).

The first logical place to look for guidance on the process of developing and tracing requirements is in published systems engineering standards such as those distributed by NASA and IEEE. However, neither the NASA Systems Engineering Handbook [3], the IEEE International Standard for Requirements Engineering [2], nor the IEEE International Standard for Systems Engineering Planning [4] address science or measurement requirements specifically. In fact, despite providing a relatively comprehensive list of requirement types, IEEE does not discuss any requirement type that really captures the unique purpose of the set of science and measurement requirements. [2]

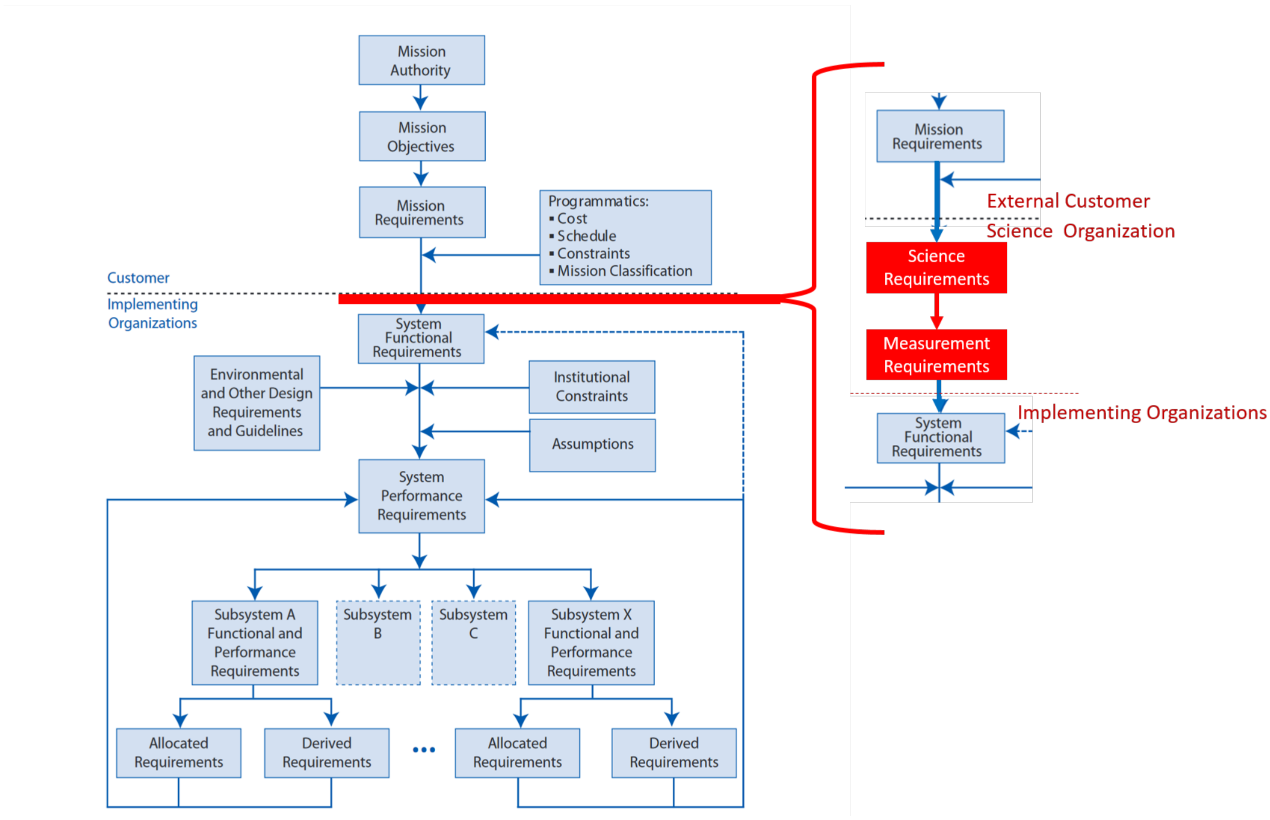

Figure 1 A modified flow down of requirements, from [3] with modifications in red to show science and measurement requirements flow down

The NASA handbook discusses the requirements flow process as shown in Figure 4.2-3 from reference [3]. In this figure, the mission objectives are translated into mission requirements (L1s using the terminology in this paper), and from there they flow directly into the development of system functional and performance requirements. The handbook describes functional requirements as specifying "what functions need to be performed," and performance requirements as specifying "how well these functions must be performed." Although this categorization of requirements is clearly applicable to specific engineering products, it does not exactly map into the language used to describe the content of the science and measurement requirements. Science and measurement requirements essentially constrain the hypotheses that are being tested in the mission, and categorize and clarify the associated range of parameters over which the experiments (and thus observations) must span to effectively test those hypotheses.

Because these requirements describe hypotheses and experiments rather than physical objects, forcing the language of "functional" and "performance" requirements to encompass these unique constraints can be non-intuitive. It is awkward (although not impossible) to describe the "function" of a dataset or the "performance" of a measurement. For example, a measurement requirement might constrain an observation to have a given the spatial resolution. While it is straightforward to write a functional or performance requirement on the instrument itself ("The instrument shall achieve a pixel scale of X", "The instrument shall collect measurements below altitudes of Y"), calling this measurement requirement a "performance" requirement is (admittedly subtly) misleading. To be precise, we would instead say that the instrument must achieve a certain performance in order to achieve an observation that has a desired quality (such as spatial resolution). Referring to the "quality" of the observation (or its "coverage" or the other descriptive fields described in this paper) is more natural than trying to abstractly reference the "performance" of the measurement.

And so while it is certainly possible to adopt the existing terminology to find ways to leverage the guidance in the handbook for science and measurement requirements, the authors would argue that such an approach encourages systems engineers to pose questions about the science and measurements in terms that are far removed from how the scientists truly think about these requirements. We naturally use slightly different language to describe an experiment with measurements of a certain quantity or quality than we do a piece hardware that collects those measurements. Although the NASA SE Handbook has the value of being very broad so as to perhaps loosely apply to almost any kind of requirement, there is still value in providing detailed, richer guidance for specific subsets of requirements that may be better understood with more specific and precise language.

So instead Figure 1 shows a modified version of the NASA requirements flow down diagram that includes the concept of science and measurement requirements as feeding into the system functional and performance requirements that can be written on traditional elements of the spacecraft. This flow down closer represents what many large science mission have implemented, and calls out the fact that requirements that constrain hypotheses and experiments perhaps require unique guidance.

Both the NASA SE Handbook [3] and the IEEE International Standard for Requirements Engineering [2] also obliquely reference different aspects to requirements writing that are relevant to the STAF. For example, IEEE provides general information on ways to elicit requirements such as brainstorming in structured workshops or studying technical documentation. [2] The suggested approach may be useful to capture the science and measurement needs, but they are so general as to not address some of the real concerns regarding science and engineering communication boundaries. The NASA handbook also describes how requirements can be partitioned into groups (based on similarities in function, performance, or "couplings"), which is a fundamental principle of the M-STAF. However, these groupings are not described beyond suggesting they can be formed by undefined "established criteria." [3] As we will describe in the next section, the STAF provides a set of defined criteria by which requirements can be grouped to aid in their development. Similarly, the NASA handbook suggests that one can develop performance requirements by writing requirements that answer questions such as: "how often and how well, to what accuracy (e.g., how good does the measurement need to be), what is the quality and quantity of the output, under what … environmental conditions, for what duration, at what range of values, at what tolerance…" [3] These questions are reminiscent of the kind of information codified into the STAF fields, as described in Section 3. Thus, the published standards provide excellent guidance on the properties of good requirements sets, and allude to the processes and mechanisms one can use to generate a good set. *However, the generality of the advice leads to some difficulty in applying all of the processes to a science-specific requirements flowdown, including the measurement requirements*.

The next source of information that can shed light onto this part of the requirements flow down is the experience of previous missions. It is difficult to find references to the process used for requirement generation on actual missions, likely because much of this information is passed on to new projects via the institutional processes (reviews, staffing cross-over, etc.). However, the use of science and measurement requirements as distinct levels used to communicate the L1 requirements into implementable engineering requirements has been referenced in literature, [5], and it is possible to find references to requirements that clearly fall into this category [6] but these papers do not focus on the process used to generate the requirements. Dodge et. al. [6] writes as a member of the Juno payload office, and he references the requirements process, but does not describe it in detail. He also notes the challenges associated with managing ten instruments across many institutions and suggests a set of programmatic solutions to ensure close communication among the teams in order to help avoid misunderstandings. As with the published standards, the available literature clearly touches on these issues and alludes to the solution, but does not provide a significant amount of detail that would be useful in implementing such a process on a new project.

One tool that has emerged as a standard for proposals is the Science Traceability Matrix, or STM. [7] This tool is essentially a two-dimensional structure where the column headings represent relevant categories and move left-to-right from most general to most specific. The rows in this structure are organized by science objectives, with subrows forming the associated measurement objectives. The matrix then calls for the matching of measurement requirements, instruments, instrument requirements, and data products to these various objectives. Organizing this information in this way allows for an individual to quickly trace between science requirements and the performance requirements that flow from them. The brilliance of the STM, however, is that it provides a structured way to communicate science and engineering relationships. The fact that this tool is widely recognized across projects and institutions (in part because it often forms part of a proposal to NASA) means that it is well-positioned to establish a common language in a context that can be understood by both the science and engineering communities.

Although Weiss, et. al. noted in his 2005 paper that there are typically "insufficient workforce resources to update the STM" after the proposal, [7] the STM has evolved such that on a few active projects, a unified STM (USTM) is created by the science team. The USTM can explain how instrument-level STMs from the instrument proposals, individually and as a payload, support the mission goals. Unfortunately, examples of these (already sparse) USTMs are not widely available in literature.

The STM is a powerful tool that can capture the traceability of some of the mission and instrument performance requirements, hence STAF builds on its basic idea and expands on it so that STAF can actually be a tool to generate a valid set of measurement requirements. STAF goes beyond the STM specifically by:

a. Being structured in such a way as to be readable and usable by both scientists and engineers,
b. Providing a structured language that facilitates communication amongst all the stakeholders of the measurement requirements, and
c. Providing a method that emphasizes completeness and consistency across measurement requirements from different instruments.

3. FRAMEWORK DESCRIPTION

It is widely recognized that communication across different institutions, cultures, and roles is a significant challenge in technical projects.[8] One way to facilitate this

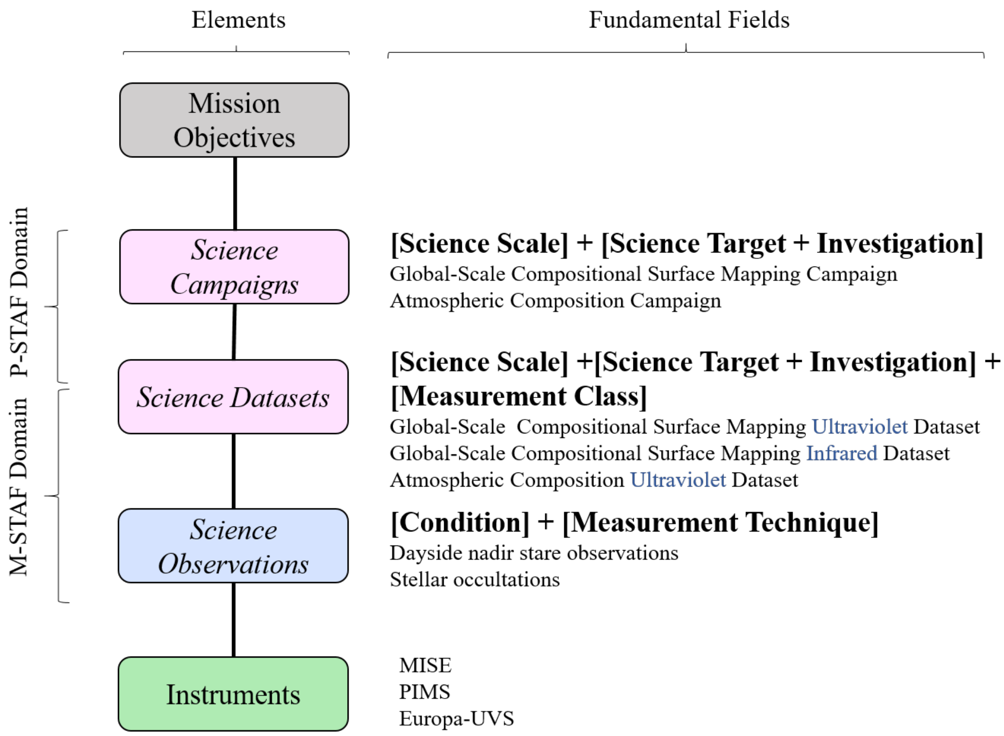

Figure 2 The taxonomy of the STAF elements

communication is to create a self-consistent language and framework that is recognized by all the stakeholders and used by all parties when in the "trading zone": a pidgin. [9] The STAF attempts to develop this taxonomy pidgin to facilitate discussions across the science-engineering divide in a specific, detailed way suitable for implementation. This consensus-building exercise results in a compromise that simplifies the science boundaries on a project in order to enable engineers to perform analyses using these basic distinctions, but improves the reporting fidelity of the engineering assessments when they provide results to the scientists, which allows for more insightful and informed conversations about how to address engineering challenges that affect science.

So what is this pidgin? STAF uses an organizational concept of "fields", or categories, that describe the information which must be captured in a requirements set. Fields define categories of knowledge that different stakeholders need in order to complete the requirements flow down. These fields also allow engineers to appropriately bin and analyze the requirements throughout the design phase.

There are two types of fields in the STAF: **fundamental fields**, which are used to create the elements of the STAF taxonomy (see Figure 2), and **descriptive fields**, which are used to describe categories of requirements associated with the system elements. The fundamental fields are prescriptive – i.e., once a master list has been developed for that field, then its designation for a given element must come from the master list. The descriptive fields, on the other hand, act as labels for different types of requirements that enable stakeholders to quickly identify the requirements that most readily apply to their needs.

The fundamental fields are **science scale**, **science target+investigation, measurement class, technique,** and **condition.** The first three are mainly used in the P-STAF domain (see [1]), while in the M-STAF domain we focus on the last two. Note that a campaign is defined by a science scale and science target+investigation, so the campaign name is embedded in the science dataset name.

The descriptive fields are **spatial coverage and distribution, temporal coverage and distribution, diversity and special cases, internal measurement correlations,** and **measurement quality.** These fields are used to develop the measurement requirements once the elements have been defined.

STAF Hierarchy of Elements

Because different stakeholders link into these requirements at different levels, we first start by defining a taxonomy in Figure 2 that hierarchically defines the key elements in the STAF: *science campaigns*, *science datasets*, and *science observations*. The customer requirements (manifestations of the mission objectives of the mission) are taken as inputs to this framework. Similarly, requirements associated with instruments (and other subsystems at the same level) are outside the purview of the STAF. However, the instrument needs to be understood as a specific, named element so it can be distinguished from the other elements in the STAF.

Table 1 M-STAF-domain key terms

Science Dataset	An organizational structure that is used to distinguish between the different science contributions of a given instrument. The science dataset is composed of [Science campaign] + [Measurement class] distinctions. For example, the Atmospheric Composition Ultraviolet Dataset
Measurement Technique	Used to distinguish between different approaches to collecting observations within a single measurement class, especially if they have a distinct set of spacecraft interfaces, coverage, quality, or condition requirements. For example, a scan is often distinguished from a simple nadir stare because it requires unique spacecraft interfaces and may have unique quality or frequency requirements.
Conditions	Used to distinguish between observations that must be collected under different geometric or configuration states. For example, a local solar time condition can distinguish between a dayside and a nightside observation or an altitude condition can distinguish between high observations or close observations.

In the P-STAF domain, the *science campaigns* can be thought of as a construct used to group together related science hypotheses and associated investigations that satisfy a given L1. For example, a mission may have a "global-scale surface mapping" campaign, or an "active plume search" campaign. More detail about the definition and subtleties of the *science campaign* element can be found in our companion paper. [1].

At a level below the *science campaigns*, STAF proposes a construct called a *science dataset*. The *science datasets* are groupings of observations collected by a single instrument or investigation that can be used to address a given *science campaign*. In other words, any data an instrument collects that can help address a given *science campaign* falls into an associated *science dataset*. However, the word dataset is yet another term that is often used in different ways by different communities, so it is important to be clear about its definition and naming. The easiest way to maintain this clarity is to name the science dataset with the name of the campaign it addresses and the generalized type of measurement being made (**measurement class**). For example, an ultraviolet imaging spectrograph performs ultraviolet imaging spectroscopy. The measurement class for that instrument, depending on the instrument team's preference for brevity, may be "ultraviolet imaging spectroscopy" or simply "ultraviolet." Thus, if the ultraviolet instrument contributes to the "active plume search campaign," it would have an associated "active plume search ultraviolet dataset." More information about the *science dataset* element can be found in our companion paper. [1]. Instead, this paper takes as given that the system has developed a *science campaign* and *science dataset* master list at the P-STAF level and instead focuses on the M-STAF domain, which describes the connection between the *science datasets* and the *science observations.*

The *science datasets* are populated by individual *observations* that collectively make up the set of data that supports specific *science campaigns*. These observations can be thought of as the individual data unit that are generated by a given *instrument*. The observations can be called images, cubes, groundtracks, or simply measurements. These terms are not constrained independently (and are therefore used interchangeably) by the STAF, although some missions may find it useful to make a distinction. We have found that different instrument teams prefer different terms and that the terms are so close in common-use language that making a distinction caused unnecessary confusion among the different stakeholders.

Fundamental Fields in the M-STAF Domain

There are five fundamental fields in the STAF: **science scale, science target+investigation, measurement class, measurement technique**, and **conditions**. For the M-STAF, the first three are codified in the science dataset element, as shown in Table 1. An example set of science datasets compiled for the Europa Mission that follow this framework can be found in the companion paper [1]. Here we concentrate instead on the two remaining fundamental fields that relate to the M-STAF domain – **measurement techniques** and **conditions** – both of which are used to define a *science observation*.

A *science observation* distinguished by the **technique** used to collect it (scanning observations are distinct from stellar occultations), and (when applicable) the **conditions** under which it must be collected (a dayside nadir observation is distinct from a nightside nadir observation). These two fields (technique and conditions) form the basis of the measurement level of the STAF.

It is important to note that observation definitions are specific to a given **measurement class.** An infrared stellar occultation is distinct from an ultraviolet stellar occultation, for example, because they may use different star catalogs and/or have different conditions that may apply. It is possible to make this distinction clear in the measurement text, or (at the risk of some redundancy in the text of the requirement) it can be formally written into the name of the observation ("ultraviolet scans" or "thermal scans", for example).

Each individual investigation team can develop their own list of **measurement techniques** appropriate to the way their science (and instrument) works. For example, the ultraviolet instrument may use stellar occultations, solar occultations, Jupiter transits, nadir observations, and scans as distinct measurement techniques. However, a visible class may

Table 2 Example of a mapping between measurement classes and their unique measurement techniques

Measurement Class	Measurement Technique		
Impact Mass Spectrometry (IMS)	Endogenic Particle		Anion
	Endogenic Particle		Cation
	Exogenic Ring Particle		(Either)
	Exogenic Nanograin Particle		(Either)
Infrared	Nadir Stare		
	Scan		
Magnetic	(none)		
Neutral Mass Spectrometry (NMS)	Ambient		
	Cryotrap		
Plasma	Ions		
	Electrons		
Radar	VHF	Sounding	
	VHF	Reflectometry	
	VHF	Altimetry	
	HF	Sounding	
	HF	Reflectometry	
	HF	Altimetry	
	VHF/HF	Plasma	
Thermal	Nadir Stare		
	Scan		
Ultraviolet	Nadir Stare		
	Scan		
	Stellar Occultation		
	Solar Occultations		
	Jupiter Transit		
	Neutral Cloud/Torus Stare		
Visible	(Surface)	Panchromatic	(Monoscopic)
	(Surface)	(Panchromatic)	Stereo
	(Surface)	Color	(Monoscopic)
	Limb Profile	Panchromatic	(Monoscopic)
	Geodesy	Panchromatic	(Monoscopic)
	Terminator	Panchromatic	(Monoscopic)
	Limb	Panchromatic	(Monoscopic)

identify the color, number of images, and specific observation target as the elements of their observation techniques ("limb panchromatic stereo images"). Examples of potential technique types are shown in Table 2, but it is important to note that no generic list can capture the techniques that are meaningful to all possible instruments across a given class. These working examples simply illustrate the diverse range of forms that a technique name might take. When considering whether to include a technique in the list, consider whether or not the investigation team needs to write requirements specifying unique qualities, coverages, or conditions on that type of observation. If so, then it should be included in this list. The technique list may also be a null set if the measurements are not distinguished in any meaningful way by the investigation team.

Conditions describe the configuration and geometry of the spacecraft and the planetary bodies when an observation is collected. There are many different types of conditions, including: lighting (local solar time, emission angle, etc.), position of the spacecraft relative to the body (altitude or latitude, etc.), and velocity of the spacecraft (relative to the body or Keplerian ram, etc.). These conditions can be defined at any relevant part of the trajectory. For example, one might apply a condition at the point of closest approach to a body or the point of the observation collection. It is important to note that information is only a condition when it must be true in order for the measurement it describes to be valid. If the team defines an observation as a low-altitude scan, then scanning measurements made above a certain altitude are no longer valid "low altitude scans." That altitude becomes a condition on the "low altitude scanning observation." If, on the other hand, a science team wants to collect a scanning measurement at a variety of altitudes, the scanning measurement condition should specify the altitude range that makes for a valid scan, and the altitude distribution and minimum number of needed scans are covered under a different field.

By combining these fields, the *observation* element can be defined by the combination of **measurement technique** and **conditions** used to collect it. The team writing requirements on this element has the ability to give any observation an appropriate moniker for clarity (for example, "low-altitude nadir stares") so long as there is a corresponding requirement or definition explaining the valid altitude range that is classified as "low-altitude." Alternatively, if a given technique always pairs with a set of conditions, the requirements authors may choose to just call the observation by its measurement technique alone (for example, if a specific instrument only ever performs a scan between 100,000 km and 30,000 km altitude with a local solar time between 3:00 am and 9 pm, using the term "scans" for that measurement class should imply these conditions apply – and further, the condition information should not need to be repeated in every requirement that constrains something about the scan.

It is also important to note that any given observation may contribute to many different science datasets. The same ultraviolet stellar occultation, for example, may contribute to both the atmospheric composition ultraviolet dataset and the active plume search ultraviolet dataset. This many-to-many mapping between datasets and observations is acceptable and expected in the STAF, and is made clear by a proposed template for writing the text of the requirement described in subsequent sections. Now with a clear concept of a *science observation*, it is possible to develop the measurement requirements.

M-STAF Descriptive Fields

Once the master lists for the fundamental fields are developed, the M-STAF proposes a set of descriptive fields that classify the information needed to be captured in the measurement requirements themselves. There are five main categories of descriptive fields: **spatial coverage and distribution, temporal coverage and distribution, diversity and special cases, internal measurement correlations,** and **measurement quality.**

The **spatial coverage and distribution** field is a category for defining the minimum acceptable number of observations, whether written as a specific discrete number ("100 stellar occultations") or a percentage coverage of a body's surface

Table 3 Example of a mapping between measurement classes and their unique measurement qualities

Measurement Class	Quality of Measurements
Impact Mass Spectrometry (IMS)	Molecular Weight Range
	Mass Resolution
	Particle Size Range
	Particle Size Accuracy
	Confidence of Particle Composition
	Spatial Resolution Quality Metric
	Accuracy of Boresight Velocity
	Particle Impact Rates
	Charge Uncertainty over Range
Infrared	Pixel Scale
	Spectral Coverage
	Spectral Contrast
	Single Image Coverage
	Smear
	Stability
Magnetic	Number of Axes
	Range
	Frequency
	Accuracy
Neutral Mass Spectrometry (NMS)	Mass Range
	Mass Resolution
	Dynamic Range
	Contamination
	Sensitivity
	Surface Sample Resolution
Plasma	Frequency
	Energy/Charge
	Energy Resolution
	Entrance Flow Angle Accuracy
Radar	Track Length
	Frequency
	Observation Depth
	Vertical Resolution
	Along Track Resolution
	Range Error
	Fast Time Resolution
	Sensitivity to Contrast Permittivity
Thermal	Spectral Bandpass and Resolution
	Spatial Resolution at Altitude
	Radiometric Accuracy over Range
	Radiometric Range and Precision
	Swath Width
	Number of Filters
	Smear
Ultraviolet	Spectral Bandpass and Resolution
	Spatial Resolution at Altitude
	Sensitivity
	Sampling
Visible	Smear
	Signal to Noise Ratio
	Pixel Scale
	# Color Bands
	Wavelength
	Arc Length
	Convergence Angle

(70% of the surface). It may include requirements on a minimum number of landforms that must be imaged by the observations (provided that the list and map of acceptable landforms is available to the mission design team). This field also includes any requirements specified on the desired distribution of observations, for example across latitude or longitude.

The **temporal coverage and distribution** field categorizes similar requirements, but focuses on those that must be specified in time rather than space. For example, it may be important to specify a minimum amount of time a given observation has, or the frequency at which the measurement is collected.

The **diversity and special case** field is not always populated, but it captures the fact that there may be unique observations that provide an important counterpoint in the science. These observations are collected using the same technique as other observations but are collected primarily so the set of data can achieve a minimum diversity needed for important scientific comparisons. For example, an investigation may want to have at least one high latitude observation, but may not want to specify other constraints (performance requirements, spatial coverage requirements, etc.) on this relatively rare and potentially difficult-to-achieve observation. Calling out these unique observations as a separate field enables systems engineers to track these more rare observations separately because they pose a different risk profile than the observations that will be repeated more frequently.

The **internal correlations field** describes requirements that need to be connected to other measurements made by the same instrument. For example, the radar instrument may want to have a certain number of groundtrack intersections to support their science. The requirements that dictate how two observations must relate to another fall into this field. Similarly, some thermal images may need to be collected of the same geographical location but at two different local solar times in order to perform a comparison. The minimum overlap of the night and day observations can be specified in this field. This field does not capture information where a given instrument needs another instrument to support its science; we pull this information into the dataset level of the STAF, as described in more detail in the [1].

Finally, the **measurement quality** field encompasses all requirements that specify characteristics of a measurement in order to support the science. For example, an imager may need to specify a spectral bandpass and resolution, or a pixel scale at an altitude. A radar, on the other hand, may want to specify vertical resolution or range error. The list of potential measurement qualities is developed by the science team implementing that particular investigation (typically an instrument team). Although these qualities are unique to each measurement class, comparing all of the qualities across the measurement classes can help catch missing requirements. For example, if an ultraviolet imager specifies its spectral bandpass and resolution but a visible imager does not, it is possible to ask the team if a requirement might be missing. In some cases, there may be a fundamental reason why the instrument does not need to specify the same qualities as other similar instruments, but that explanation is more meaningful and allows a systems engineer to ask more informed and insightful questions. Table 3 shows an example mapping of measurement class and possible measurement qualities. Note that this list is by no means exhaustive and can vary from one science investigation team to the next, depending on the needs of the mission.

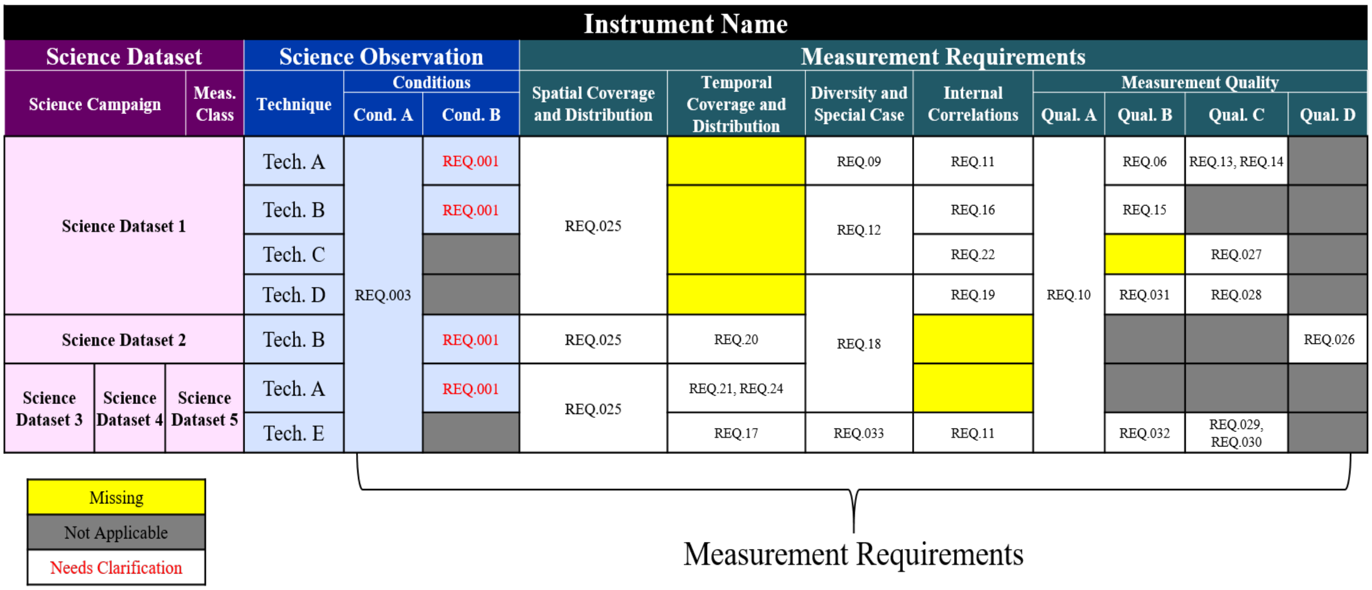

Figure 3 M-STAF Matrix generic example

These five descriptive fields enable scientists and engineers to sort requirements in a meaningful and consistent way. At the measurement level, requirements are written on observations distinguished by their **science dataset**, **measurement technique**, and **conditions**, and those requirements should fall into one of these descriptive fields. These fields provide a way of structuring the conversation between the systems engineers and the scientists to ensure that the requirements set is not missing information.

M-STAF Matrix

The true value of the taxonomy of the M-STAF becomes apparent when it is applied to a set of tools that the project can use to increase efficiency and consistency. The M-STAF matrix, shown as a generic example in Figure 3, is one such powerful tool. It provides a visual representation of the measurement requirements which can be used to quickly communicate a large amount of information while simultaneously checking for completeness across the datasets. It preserves the STM advantage of enabling a team to see how specific measurements contribute to different science, but it uses the construct of a *science dataset* to group the science rather than grouping by science objectives as the STM does. Perhaps most importantly, this matrix can be interpreted by scientists, engineers, and eventually even computers, which bridges the communication gap and reduces the tedious work of attempting to classify and link every requirement by hand in order to merely start the conversation.

It is perhaps easiest to start by constructing a single M-STAF matrix for a given measurement class. This organization best complements the instrument-specific team organizational structure that often already exists within the science team. Other matrices that leverage the STAF can be constructed to support different stakeholder needs (for example, a matrix that summarizes all of the conditions for all science observations may be valuable to the trajectory design and mission planning team), but this M-STAF matrix is most useful for the construction and parsing of the measurement requirements themselves.

In the M-STAF matrix, both the fundamental fields and the descriptive fields are arranged as column headers in the matrix, as shown in Figure 3, where fundamental fields are on the left and descriptive fields are on the right. The *science dataset* names are written as merged row headings, while each *science observation* that supports a given dataset is written as a sub-row in the matrix. Thus, every row represents a different observation that contributes to a given *science dataset*. From this basic structure, the matrix is populated with measurement requirements, usually written in shorthand with a link to the proper text of the requirement itself. The matrix can be used as a structure for generating those requirements, or it can be used to sort and check the requirements once they have been written. In either case, when it is not clear if a requirement applies to a specific cell in the M-STAF matrix, it likely needs to be reworded to be clear over which science dataset and science observation it applies. Similarly, if there is no requirement that specifies that particular descriptive field for a given science dataset and observation combination, it should be highlighted as a potential missing requirement. In some cases, those cells will not be applicable to a given measurement class, dataset, or observation. In that case, the cell is greyed out and ignored. *However, the fact that the matrix triggered the conversation ensures that these cells were intentionally not captured in the requirements rather than accidentally left out of the set.*

Each cell of the M-STAF matrix can include the requirement's essence written in a short-hand format. However, one could imagine that with the advent of model-based systems engineering approaches, this may become a view in the system model that links to the requirement text itself. We also added the requirement number to allow for easy cross-reference between the requirements list (more conventional way to maintain the requirements) and the M-STAF matrix.

Once the M-STAF matrix is constructed, it serves as a table-of-contents for the measurement requirements themselves (rather than scrolling through a list, each cell identifies the number of the appropriate requirement). It can be used to directly compare datasets and make it clear which observations affect which type of science. Similarly, it can be used to develop similar language for requirements that are in the same descriptive field. This representation has a number of uses for many different stakeholders in this process, which are described in detail in the following sections.

4. Framework Implementation

As members of the Europa Mission payload team, the authors have the opportunity to use this framework to categorize measurement requirements for a diverse set of instruments selected for the planned Europa Mission. The framework has been specifically developed to be as instrument-agnostic Mas possible. For this paper, we use two instruments on the Europa Mission to provide practical example of the M-STAF implementation: Europa-UVS (Europa Ultraviolet Spectrograph), led by Kurt Retherford at Southwest Research Institute) and PIMS (Plasma Instrument for Magnetic Sounding), led by Joe Westlake at John Hopkins Applied Physics Laboratory.

How to Build the M-STAF Matrix

As one can see from the M-STAF matrix template, in order to start populate the matrix for any measurement class, the list of its science datasets must be available and settled, as described in the P-STAF domain. [1] Once that step is completed, we start by compiling the lists for the M-STAF specific fundamental fields **measurement technique** and **conditions**. Europa-UVS, for example, has identified six different measurement techniques (listed in the fourth column of its M-STAF matrix in Figure 4). These techniques are distinguished in a few ways. In some cases, the fact that they use notably different methods of observing a phenomenon is important (a stellar occultation observes how the light from a bright UV star is absorbed as it passes through an atmosphere, whereas a nadir observation observes reflected UV light from the surface of a body, for example). In other cases, the techniques are distinguished by the fact that they use significantly different sources and/or elements on the instrument (for example, a solar occultation involves pointing the instrument's solar port toward the sun, whereas a stellar occultation involves pointing the instrument's airglow port to a bright star, even though both observations use UVS absorptions to make the measurement).

Measurement techniques in this context might not indicate something that the instrument does differently at all but instead differentiates the observations in ways that must be called out separately in the measurement requirements. If the team needs to constrain the observations differently, such as their coverage or quality, it is a good indication that they represent different techniques. For example, PIMS, a plasma instrument, collects and measures properties of the plasma which consists of charged particles both positive and negative. PIMS has to specify how well the instrument needs to determine the entrance angle of the ions, but not of the electrons. Similarly, the energy range PIMS needs to measure is different for differently charged particles (see columns energy/charge and accuracy of entrance flow angle in Figure 5). Thus, ions and electrons are two different plasma measurement techniques.

Distinguishing the conditions for the different science observations follows a similar pattern, and these two instruments provide another study in contrasting instrument needs that both fit into the framework. For example, lighting conditions such as sun phase angle, incidence, and emission angles typically have to be constrained for imaging instruments. For Europa-UVS, a specific range of local solar times (codified in requirement UVS.026 and UVS.027, as seen in Figure 4) distinguish between nightside and dayside nadir observations. The team needs to specify different coverage and spatial resolution requirements for these different types of observations. Thus, the conditions field provides a way to identify the sun-Europa-probe geometry Europa-UVS needs for each of these science observations. On the other hand, the spacecraft speed is important for PIMS: if the speed is too low, it cannot capture enough particles of Europa's ionosphere. Similarly, just as light conditions are almost always important to imagers, the angle between instrument boresight and key directions like Keplerian ram or spacecraft ram are important to particle and plasma instruments. Thus, as can be seen in the PIMS M-STAF matrix, the boresight orientation is listed as a condition (see the conditions columns in Figure 5). Despite how different these two instruments are, there are similarities in the conditions they need to specify. For example, both Europa-UVS and PIMS specify the acceptable altitude ranges for the observations.

One of the valuable aspects of the M-STAF matrix is that it is a way to visually show that the conditions requirements must be applied across all of the other measurement requirements that invoke a specific observation. In the matrix, a set of conditions must apply over the entire row. This helps, for example, a mission designer to ensure that the spatial coverage of a given observation is only assessed when these conditions are true. Similarly, the instrument teams may assume that these conditions are valid when constraining their measurement quality for given observation.

Even more fundamentally, one of the most significant findings of the M-STAF development process was the power of this framework to achieve cross-pollination: it is easy to share and compare information not only amongst similar instruments but also across very diverse instruments because the fields are instrument-agnostic. Developing the M-STAF matrices for each instrument in the payload ultimately made the requirements set for each instrument stronger and more

Europa-UVS															
Science Dataset			Science Observation				Measurement Requirements								
				Conditions							Measurement Quality				
Science Campaign		Meas. Class	Technique	Europa Solar Phase Angle	Jupiter Solar Phase Angle	Altitude @ Meas.	Spatial Coverage and Distribution	Temporal Coverage and Distribution	Diversity and Special Case	Internal Correlations	Spectral Bandpass and Resolution	Spatial Resolution at Altitude	Scale Height Resolution	Sensitivity	Sampling
Global-Scale Compositional Surface Mapping		Ultraviolet	Nadir Stares	Day < 90 deg (UVS.026)		< 30,000 km (UVS.036)	70% of surface (UVS.001)	Acquisition over duration of nadir subphase (UVS.018)			<=6 nm btwn at least 150-180 nm ; <=25 nm btwn at least 105-180 nm (UVS.003: UVS.004)	<= 30 km per pixel @30,000 km (UVS.002)		TBR SNR >= 3 per spatial resolution element given a Lyman-alpha albedo of 1% at	Capable of Nyquist sampling (UVS.033)
			Scans	Day < 90 deg (UVS.026)		< 30,000 km (UVS.037)									
			Nadir Stares	Night >= 90 deg (UVS.027)		< 36,000 km (UVS.038)	10% of surface across >= 5 representative regions (UVS.009)	Acquisition over duration of nadir subphase (UVS.018)				<= 100 km per pixel @ 36,000 km (UVS.008)		TBR SNR >= 3 per spatial resolution element given a Lyman-alpha albedo of 1% at 2 deg phase (UVS.034)	
			Scans	Night >= 90 deg (UVS.027)		< 30,000 km (UVS.037)									
Landform Composition		Ultraviolet	Nadir Stares	Day < 90 deg (UVS.026)		< 360 km (UVS.039)	>= 30 representative landforms in >= 11 Europa Panels (UVS.011)		>=1 image with Europa latitude @ CA > 45 deg (UVS.012)		<= 6 nm btwn at least 150-180 nm ; <= 25 nm btwn at least 105-180 nm (UVS.003: UVS.004)	<= 1 km per pixel @ 360 km (UVS.010)		TBR SNR >= 3 per spatial resolution element given a Lyman-alpha albedo of 1% at 2 deg phase (UVS.034)	Capable of Nyquist sampling (UVS.033)
Atmospheric Compositio n	Space Environment Composition	Ultraviolet	Nadir Stares			< 390,000 km (UVS.030)	1 image in each Europa Panel and each local solar time bin (UVS.017)	>= 1 image in each combo of Europa Panel and local solar time bin; acquisition over duration of nadir subphase (UVS.017; UVS.018)			<= 2 nm btwn at least 60-180 nm (UVS.005)	TBR [30] km per pixel @ TBR [30,000] km (UVS.024)			Capable of Nyquist sampling (UVS.033)
			Scans			< 390,000 km (UVS.030)	>= 6 per flyby, distributed evenly on inbound and outbound (UVS.019)	<= 2 hours apart, distributed over >= 6 hours; distributed over >= TBR [18] months (UVS.021, UVS.031)				TBR [500] km per pixel @ TBR [165,000] km (UVS.023)		TBR SNR >= 3 per spatial resolution element given an emission brightness of 0.1 Rayleighs near 130 nm (UVS.035)	
			Stellar Occ				>= TBR [100] with at least 1 in every Europa Panel (UVS.014)						<= 50 km (UVS.013)		Continuous sampling from 400 km to Europa surface
			Solar Occ				>= 1 (UVS.015)								
			Jupiter Transit		< 120 deg (UVS.029)	< 350,000 km (UVS.028)	>= 10 (UVS.016)								
			Neutral Cloud and Torus Stare			>= 500,000 km (UVS.032)	>= 1 per orbit for >= 20 orbits (UVS.022)								
ActivePlume Search		Ultraviolet	Nadir Stares			< 390,000 km (UVS.030)	1 image in each Europa Panel and each local solar time bin (UVS.017)	>= 1 image in each combo of Europa Panel and local solar time bin; acquisition over duration of nadir subphase (UVS.017; UVS.018)			<= 2 nm btwn at least 100-140 nm (UVS.006)	TBR [30] km per pixel @ TBR [30,000] km (UVS.024)			Capable of Nyquist sampling (UVS.033)
			Scans			< 390,000 km (UVS.030)	>= 6 per flyby, distributed evenly on inbound and outbound (UVS.019)	<= 2 hours apart, distributed over >= 6 hours; distributed over >= TBR [18] months (UVS.021, UVS.031)				TBR [500] km per pixel @ TBR [165,000] km (UVS.023)			
			Stellar Occ				>= TBR [100] with at least 1 in every Europa Panel (UVS.014)						<= 30 km (UVS.020)		Continuous sampling from 400 km to Europa surface
			Solar Occ				>= 1 (UVS.015)								
			Jupiter Transit		< 120 deg (UVS.029)	< 350,000 km (UVS.028)	>= 10 (UVS.016)								

Figure 4 An example of an M-STAF Matrix using example Europa-UVS measurement requirements

complete.

Once the *science datasets*, **measurement techniques**, and **conditions** lists are compiled and used to set up the left-hand side of the matrix, we can move to the rest of the M-STAF descriptive fields. Note that each row in the matrix is a *science observation*, so by moving from left to right in the M-STAF matrix, the scientists can fill out the cells (i.e. identify requirements) for a given observation in support of a specific science campaign. Engineers can similarly use the matrix to quickly identify which requirements contain the information they need, and ensure that they understand how the observations contribute to the science datasets.

<table>
<tr><th colspan="17">PIMS</th></tr>
<tr><th colspan="5">Science Dataset</th><th colspan="4">Science Observation</th><th colspan="8">Measurement Requirements</th></tr>
<tr><th colspan="4" rowspan="3">Science Campaign</th><th rowspan="3">Meas. Class</th><th rowspan="3">Technique</th><th colspan="3">Conditions</th><th colspan="3" rowspan="3">Spatial Coverage and Distribution</th><th>Temporal Coverage and Distribution</th><th colspan="4">Measurement Quality</th></tr>
<tr><th rowspan="2">Altitude @ Meas.</th><th rowspan="2">S/C Speed</th><th rowspan="2">Angle btwn Keplerian Ram and Boresight (of a cup)</th><th rowspan="2">Local Solar Time</th><th rowspan="2">Frequency</th><th rowspan="2">Energy per Charge</th><th rowspan="2">Energy Resolution</th><th rowspan="2">Accuracy of Entrace Flow Angle</th></tr>
<tr></tr>
<tr><td rowspan="6">Ocean Properties</td><td rowspan="6">Iceshell Properties</td><td rowspan="6">Inferred Plume Evidence</td><td rowspan="4">Space Environment Composition</td><td rowspan="6">Plasma</td><td>Ions</td><td rowspan="2">Survey mode: altitude >TBR [66,000] km (PIM.001)</td><td rowspan="2"></td><td rowspan="6"><45 deg (PIM.013)</td><td colspan="3" rowspan="2">Coverage > 50% of each petal in space , evenly along petal (PIM.015)</td><td></td><td rowspan="2">>= TBR [0.02] Hz (PIM.019)</td><td>TBR [0.02] keV - TBR [6] keV (PIM.021)</td><td rowspan="6">Better than TBR [DE/E=15%] (PIM.022)</td><td>Better than or equal to TBR [5] degrees, max (PIM.023)</td></tr>
<tr><td>Electrons</td><td></td><td>TBR [0.01] keV - TBR [2] keV (PIM.031)</td><td></td></tr>
<tr><td>Ions</td><td rowspan="2">Magnetospheric mode: altitude TBR [66,000]- TBR [13,000] km (PIM.003)</td><td rowspan="2"></td><td></td><td></td><td></td><td></td><td rowspan="2">>= TBR [0.25] Hz (PIM.018)</td><td>TBR [0.02] keV - TBR [6] keV (PIM.021)</td><td>Better than or equal to TBR [5] degrees, max (PIM.023)</td></tr>
<tr><td>Electrons</td><td></td><td></td><td></td><td></td><td>TBR [0.01] keV - TBR [2] keV (PIM.031)</td><td></td></tr>
<tr><td rowspan="2">Atmospheric Composition</td><td>Ions</td><td rowspan="2">Transition mode: altitude TBR [13,000] - TBR [2,300] km (PIM.004)</td><td rowspan="2">>3km/s (PIM.006)</td><td></td><td></td><td></td><td></td><td rowspan="2">>=1 magnetospehric mode and 1 ionospheric mode measurements with freq TBR[0.2] or higher (PIM.017)</td><td>TBR [1] eV/q - TBR [50] eV/q (PIM.020) and TBR [0.02] keV - TBR [6] keV (PIM.021)</td><td>Better than or equal to TBR [5] degrees, max (PIM.023)</td></tr>
<tr><td>Electrons</td><td></td><td></td><td></td><td></td><td>TBR [1] eV/q - TBR [50] eV/q (PIM.020) and TBR [0.01] keV - TBR [2] keV (PIM.031)</td><td></td></tr>
<tr><td colspan="2" rowspan="6">Inferred Plume Evidence</td><td colspan="2" rowspan="6">Atmospheric Composition</td><td rowspan="6">Plasma</td><td>Ions</td><td rowspan="6">Ionospheric mode: altitude <TBR [2,300] km (PIM.005)</td><td rowspan="6">>3 km/s (PIM.006)</td><td rowspan="6"><45 deg (PIM.013)</td><td colspan="3" rowspan="2">>=TBR [12] flybys with CA<TBR [30] km (PIM.007)</td><td rowspan="6">Distribution intervals of CA < TBR [3] hours (PIM.011)</td><td rowspan="6">>= TBR [1] Hz (PIM.016)</td><td rowspan="6">TBR [1] eV/q - TBR [50] eV/q (PIM.020)</td><td rowspan="6">Better than TBR [DE/E=15%] (PIM.022)</td><td>Better than or equal to TBR [5] degrees, max (PIM.023)</td></tr>
<tr><td>Electrons</td><td></td></tr>
<tr><td>Ions</td><td colspan="3" rowspan="2">> TBR [11] polar flybys (<TBR[20] deg latitude from either pole) with CA <TBR [100] km (PIM.028, PIM.012, PIM.025)</td><td>Better than or equal to TBR [5] degrees, max (PIM.023)</td></tr>
<tr><td>Electrons</td><td></td></tr>
<tr><td>Ions</td><td colspan="3" rowspan="2">> TBR [15] equatorial flybys (<TBR[20] deg latitude from equator)with CA <TBR [100] km (PIM.024, PIM.027, PIM.029)</td><td>Better than or equal to TBR [5] degrees, max (PIM.023)</td></tr>
<tr><td>Electrons</td><td></td></tr>
<tr><td colspan="2" rowspan="6">Ocean Properties</td><td colspan="2" rowspan="6">Iceshell Properties</td><td rowspan="6">Plasma</td><td>Ions</td><td rowspan="6">Ionospheric mode: altitude <TBR [2,300] km (PIM.005)</td><td rowspan="6">>3 km/s (PIM.006)</td><td rowspan="6"><45 deg (PIM.013)</td><td rowspan="2">>=TBR [12] flybys with CA<TBR [30] km (PIM.007)</td><td rowspan="6">Europa true anomaly distribution intervals <= TBR [45] degrees (PIM.008)</td><td rowspan="6">The Jupiter System III Magnetic longitude distribution intervals <= TBR [45] deg (PIM.009)</td><td rowspan="6">Distribution intervals of CA < TBR [3] hours (PIM.011)</td><td rowspan="6">>= TBR [1] Hz (PIM.016)</td><td rowspan="6">TBR [1] eV/q - TBR [50] eV/q (PIM.020)</td><td rowspan="6">Better than TBR [DE/E=15%] (PIM.022)</td><td>Better than or equal to TBR [5] degrees, max (PIM.023)</td></tr>
<tr><td>Electrons</td><td></td></tr>
<tr><td>Ions</td><td rowspan="2">> TBR [11] polar flybys (<TBR[20] deg latitude from either pole) with CA <TBR [100] km (PIM.028, PIM.012, PIM.025)</td><td>Better than or equal to TBR [5] degrees, max (PIM.023)</td></tr>
<tr><td>Electrons</td><td></td></tr>
<tr><td>Ions</td><td rowspan="2">> TBR [15] equatorial flybys (<TBR [20] deg latitude from equator)with CA <TBR [100] km (PIM.024, PIM.027, PIM.029)</td><td>Better than or equal to TBR [5] degrees, max (PIM.023)</td></tr>
<tr><td>Electrons</td><td></td></tr>
</table>

Figure 5 An example of an M-STAF Matrix using a selection of example PIMS measurement requirements. Note that the Diversity/Special Case and Internal Correlation columns have been hidden because they were not applicable to the PIMS measurement requirements

It is important to stress again that the M-STAF fields provide an invaluable input to "seed" structured discussions amongst the various measurement requirement stakeholders. In filling out the quality field of an instrument that takes images, for example, the engineers can start teasing out requirements by asking the scientists if a spectral bandpass, signal-to-noise ratio, pixel scale, or smear is important for that observation. While talking to fields & particle instruments, typically the bandpass is replaced by energy range, pixel scale by energy resolution, and so forth. Because the fields are so universal, it is easy to see how once the M-STAF matrices are available for one or two instruments (no matter how much they differ), they offer a very good starting point to start and ease the conversation with scientists involved with other instruments.

As previously stated, a science observation is fully identified by measurement technique and conditions, and the same observation might contribute to multiple *science datasets*. Two cases might occur after filling out the matrix:

a) All the descriptive fields for that type of observation are identically constrained for different datasets,
b) The descriptive fields for that type of observation are differently constrained depending on the science dataset it contributes to.

Both cases are easily accommodated by the M-STAF matrix.

PIMS is an example of the first case, where all descriptive fields are identically constrained. PIMS continuously collects ion and electron measurements during the science phase of the mission. At different distances from the Europa and at different altitudes from the surface, the target quality measurement is different, but all the measurements are needed to understand ocean properties, iceshell properties, plasma composition, and to search for plumes. The fact that the observations are not strongly differentiated by science dataset can be shown in the M-STAF by listing all the datasets together on the left side of the M-STAF. This kind of M-STAF structure identifies instruments where any loss of data will affect all of the science where that instrument makes a contribution.

Europa-UVS shows an example of the second case. For the atmospheric composition ultraviolet dataset and for the plume search ultraviolet dataset, the nadir stare observations have identical coverage needs, but the measurement qualities are actually different. For example, because the atmospheric composition and active plume search science are different, they require different a spectral bandpass and resolution (in Figure 4 it is clear that the active plume search ultraviolet dataset requires a subset of the band needed for atmospheric composition science). Although clearly the instrument will be designed to cover the entire bandpass necessary for all of the supported science datasets, this information is very useful to capture. The implications here are multi-fold: a) it illuminates which science dataset is driving the instrument design, b) if there is any risk that the detector performance might not satisfy the driving requirement, M-STAF can help the instrument make a stronger case to the rest of the project about which science (and thus which customer requirements) would be impacted, c) if the instrument detector is later damaged (because of a radiation fault, for example), M-STAF can again provide some help in quantifying the science impact. In other words, M-STAF empowers the project team to crisply understand traceability as a by-product of its fields and matrix format.

How to Start Writing the Measurement Requirements

Once the M-STAF is completed, each non-empty cell is then expanded and verbalized in a formal requirement – and it is this requirement that is the governing location of the necessary constraints. Leveraging the M-STAF fields structure, it is possible to develop an example template for each type of measurement requirement (as codified by the descriptive fields). The templates are a tool that can be used as a starting point to encourage the requirements authors to use consistent language and phrasing across instruments. They are not intended to cover every circumstance, and above all the readability and clarity of the requirement text must trump any formulaic template. However, having a consistent starting point for the wording can make the process of writing them much more efficient, besides making them much easier to read as a set later in the project. Using the STAF hierarchy of elements generates the following definitions:

a) [Science Dataset] = [Science Campaign] + [Measurement Class]

b) [Science Observation] = [Condition Moniker] + [Measurement Technique]

Here the condition moniker is a shorthand way to refer to the condition distinction if applicable (nightside observations or low-altitude scans, for example). Below we have included some examples of templates and relative requirements as written for either Europa-UVS or PIMS.

<u>*Example Condition Requirement Template*</u>

For the [*Science Dataset(s)*], the [*Science Observation*] shall occur when the [**Condition** Type] is… [**Condition** Value].

UVS.026 in Figure 4 can thus be written as follows:
For all ***ultraviolet datasets*****,** all ***dayside nadir stares*** shall occur when the ***Europa solar phase angle*** is less than ***90 degrees***.

<u>*Example Coverage Requirement Template*</u>

For the [*Science Dataset(s)*], the [*Science Observation*] shall be distributed across…[Coverage value].

UVS.001 in Figure 4 can thus be written as follows:
For the ***global-scale compositional surface mapping ultraviolet dataset***, the ***dayside nadir stare*** observations shall be distributed across ***70% of the surface*** of Europa.

For the [*Science Dataset(s)*], the minimum number of

[*Science Observation*] shall be…[Coverage value].

UVS.014 in Figure 4 can thus be written as follows:
For the ***atmospheric composition ultraviolet dataset*** and ***active plume search ultraviolet dataset***, the number of acquired ***stellar occultation*** observations shall be ***greater than or equal to 100***.

Example Measurement Quality Requirement Template

For the [*Science Dataset(s)*], the [Measurement Quality Type] for the [*Science Observations*] shall… [Quality Value].

PIM.016 in Figure 5 can thus be written as follows:
For all ***plasma datasets***, the ***sampling frequency*** of the ***ions and electrons*** observations in the ***ionospheric mode*** shall be ***greater than or equal to 1 Hz***.

The key to the success of this approach is adopting a reasonable flexibility to ensure that the language is readable and clear. For example, the PIM.016 requirement uses the term "ions and electron observations in the ionospheric mode" to describe its science observations under a given altitude condition. The text was written to be sensible to both the instrument and make it clear to the engineer which observation was being constrained. Shorthand terms (for "all" plasma datasets, for example), may make the text easier to parse, as long as they are well defined.

5. Framework Functions And Uses

Decomposition Archetypes

Especially because of its tabular format, the M-STAF matrix shows traceability of a requirement up-flow to a particular dataset but traceability can also be ensured down-flow. Because all the requirements are effectively categorized by the descriptive field they belong to, after generating all the measurement requirements, it is possible to identify decomposition archetypes. The archetypes are key in ensuring that all the stakeholders for a particular requirement are informed of its existence and can work together to meet it. This approach is a powerful tool when coupled with Model Based System Engineering (MBSE). Each measurement requirement, once entered in the system model, can be tagged with the appropriate M-STAF field, and according to that, the model itself can ensure that this requirement flows to the appropriate project system/subsystem. For the Europa Mission some examples of archetypes are shown in Figure 7 (note that Figure 7 is not exhaustive). Typically the requirements in the condition, spatial and temporal coverage, diversity, and correlation fields are satisfied by mission planning (trajectory design, the mission plan development, etc.) directly. On the other hand, the measurement quality field has a much more diverse set of potential decompositions. Some qualities, such as pixel size, are entirely satisfied by the instrument design alone, others, such as smear, involve also other system such as the spacecraft, the trajectory navigation system, and the mission operations system (MOS). The existence of the M-STAF fields and the categorizations of the measurement requirements enable systems engineers to focus on developing appropriate archetypes for their mission, rather than sorting through potentially hundreds of requirements to identify suitable decompositions for each one individually. Clearly, a cognizant person must be in the loop to catch any decomposition inconsistencies or errors, but this initial guess at an appropriate decomposition significantly improves the requirement flow efficiency.

Ensuring uniform and consistent discretization of the information embedded in these requirements enables them to become machine-readable. In essence, organizing the information in a structured way within requirements, and not just across requirements, facilitates its capture in queryable information management systems. The concept of capturing information in a way that it can be queried is obviously not a novel one. What is perhaps novel is the concept of capturing the information encoded within requirements in a way that can be queried and analyzed. The 'database' (used here colloquially) could take the form of any number of currently employed requirements information capture environments, such as DOORS or SysML. The purpose is to allow the

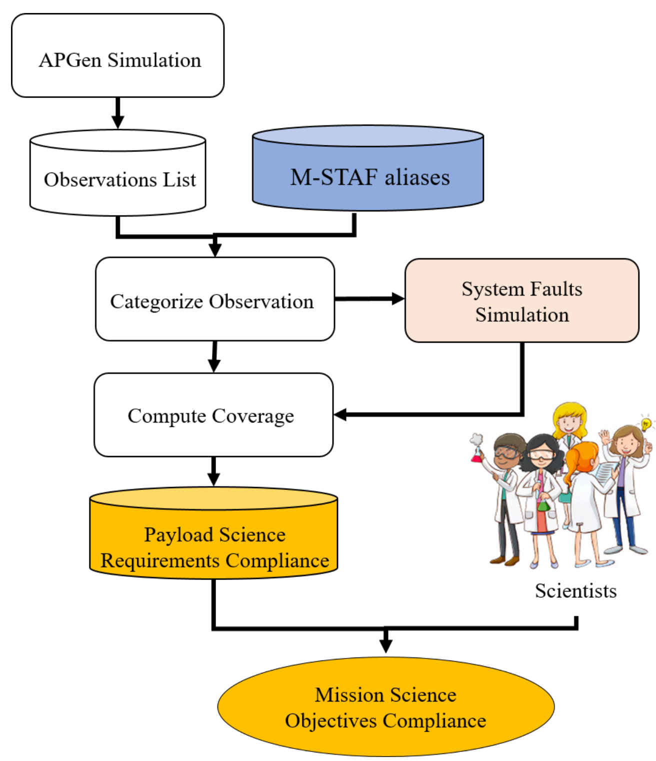

Figure 6 Mission planning team analysis block diagram

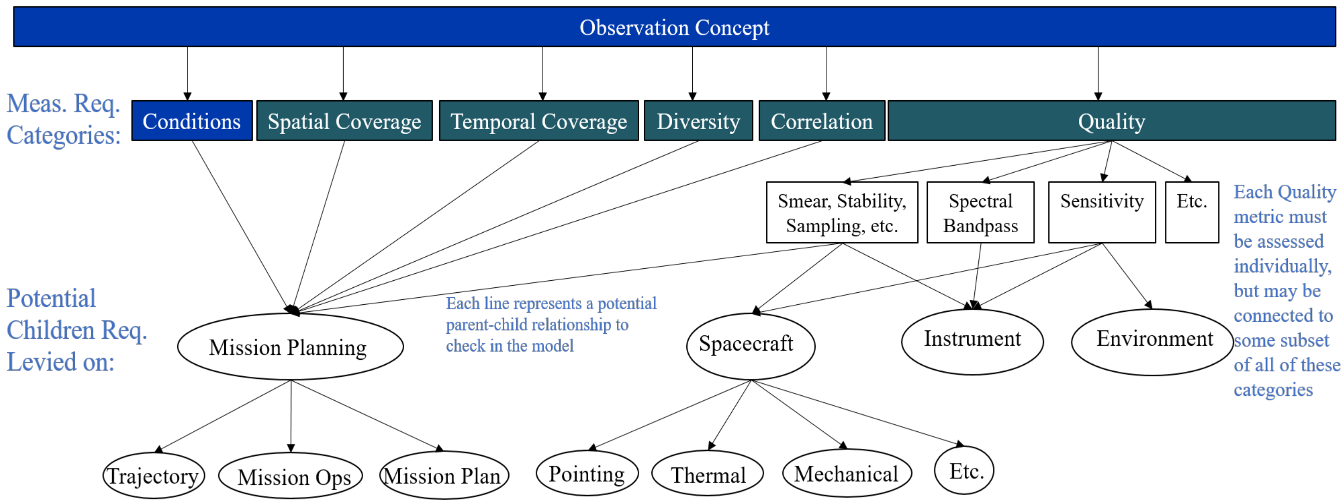

Figure 7 Decomposition archetypes examples

user to ask questions about the information in the requirements: a mission designer could access a view which shows all of the lighting constraints on trajectory designs; a scientist could look at a different view showing the overlaps and gaps in the electromagnetic spectrum coverage; a pointing systems engineer might want to know the sensitivity to smear of the optical instruments.

M-STAF Relevance to Requirement Compliance

Another potential use of the M-STAF matrices is in helping quantify the mission design performance and inform trades at the mission level. The M-STAF approach is different from other proposed value assessments [10], [11] because it takes as given a set of mission science objectives and, instead, seeks to understand the impact of engineering decisions on those objectives. In fact, many of the coverage requirements, both spatial and or temporal, cannot be verified exclusively by the instruments teams because it requires input from the mission planning team. As part of its scope of work, the mission planning team generates timeline of observations for each instrument using APGen (JPL-developed software) for each mission tour that is designed (see Figure 6). Given a specific tour, the observations timelines are the result of an optimization that takes into account a large number of constraints (including the measurement requirements) on resources such as: spacecraft power, thrusters usage, hours of reaction wheels usage, etc. Without the M-STAF in place, the mission design team must painstakingly go through each requirement and use their best judgment to build a very similar mapping – which conditions apply to which observations in the timeline – and then individually solicit feedback from individual instruments to assess the tour's compliance with the measurement requirements. Now, requirements authors that own the M-STAF matrix representation can apriori confirm that a requirement is intended to be a certain type (such as a condition, coverage, etc), and how they link together. By seeding the initial conversation (which is still necessary to ensure understanding), the task is significantly simplified. By given the requirements authors an opportunity to provide an interpretation of those requirements in advance, the requirements are more clearly communicated to other stakeholders who need to implement them.

The mission planning team on the Europa Mission conducted a case study using the M-STAF to demonstrate its value. The team was able to build an automated routine that can check compliance of a specific mission tour against the measurement requirements. As shown in Figure 8, the mission planning team can use the M-STAF (and the matrix) to create an alias for each observation using the data set name, the technique, and the conditions. Then (see Figure 6), the list of observations is compared to the aliases and categorized as belonging to the appropriate dataset. Under the assumption that no faults occurred (nominal case), it is possible to check the coverage of the datasets against the appropriate temporal/spatial coverage requirements and record the findings in the M-STAF itself, see Figure 8.

This assessment tool is able to not only identify individual requirements at risk in a given tour, but can also report back on the specific science datasets and science campaigns that are impacted as a result of a specific requirement not being met. This feedback loop enables the mission planners to improve the chance of a given tour meeting the science needs when it is being designed. It also enables the team to perform analyses on science reliability, for example, at a much more insightful level. For example, when the mission designers report back to the project science group on the impacts of radiation fault analyses, it is much more accessible to the broader community describe how the "atmospheric composition ultraviolet" dataset is affected rather than explaining how UVS.014 is not met because stellar occultations are affected.

Similarly, this approach more readily provides information on the L1 compliance, and can help system-level implementers understand how to best proceed. So, once the mission planners have an assessment tool, it is possible to have a partial assessment of the L1 compliance for the examined tour and scheduled activities. *This automated tool should by no means be used as a substitute for the scientists' judgment.* However, its results can improve efficiency by indicating to the scientists where to focus the tour evaluation efforts. For example: for a given tour, if coverage requirements for many science datasets, within a science campaign are not met, it would be advisable for the scientists to spend time at closely evaluating those datasets. This input can help them decide if a new tour should be designed to improve the science return for that particular campaign, or if the requirements are just too hard to satisfy. Because a mission (such as the Europa Mission) may develop dozens of tours over the course of the project, and their evaluation requires a lot of effort, making the evaluation as efficient as possible will help everyone across the project focus on getting as much science from the mission as possible.

The same analysis pipeline can be used to understand sensitivity of the tour to faults caused by the radiation environment and inform requirements on maximum time to recover from faults (for example). Once each observation is assigned to the appropriate dataset(s), faults are simulated by simply dropping observations at a rate and duration as prescribed by the radiation model. The output of this step can then travel through the same pipeline as the nominal case (see Figure 6) and the impact on datasets, campaigns, and partially on the L1s can be assessed. Again, this by no means would substitute the scientists' judgment. But that judgment can be much more informed if given highly relevant information, such as how robust the design is to radiation. These assessments may help develop system requirements that are otherwise difficult to determine, such as a maximum recovery time after a fault.

Ultimately, the project (including the instrument teams) must collectively decide how relevant these tools are to the verification of requirements, the design of tours, and the development of a science reliability strategy. But the value of the STAF is that it provides the project with the information and language necessary to make these decisions. The STAF is merely a tool that provides a richer structure and deeper understanding of the science flowdown in requirements space.

6. CONCLUSION

In large space science missions involving multiple instruments, it is important to have a set of requirements that clearly defines the science needs. These requirements ensure the system design is properly constrained when they are properly decomposed and interpretable to both the scientists and engineers who need to interact with them. The current state of practice lacks clear and detailed information on how to handle the unique science-driven requirements flowdown, meaning that most projects end up developing their own systems with varying degrees of success. Although the Science Traceability Matrix is an important step towards building this requirements set, it does not address the needs of all of the stakeholders in the measurement requirements.

The Science Traceability and Alignment Framework (STAF) expands on these starting points and provides a taxonomy and toolset that, if implemented, allows the system engineering team and the project science team to derive a set of measurement requirements that is complete, traceable, and consistent. This framework represents a comprehensive attempt to structure the dialog between scientists and engineers to generate requirements that can cross barriers among all their stakeholders.

Because the science requirements flowdown spans many

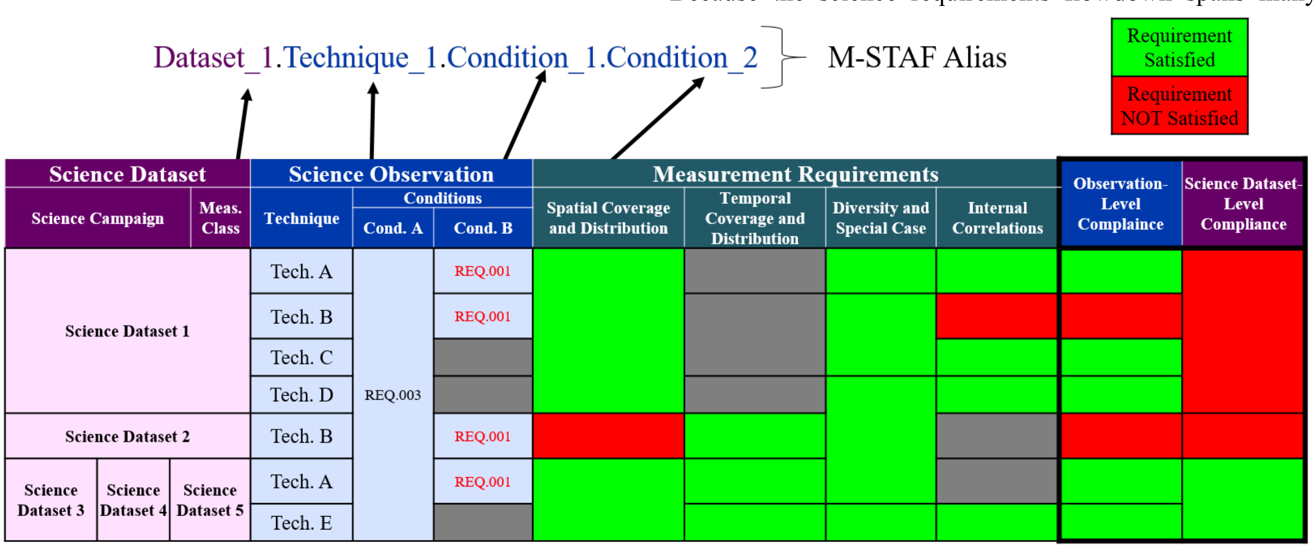

Figure 8 Measurement alias definition and measurement requirement compliance

levels of the project, the STAF is split into two domains: the project domain (P-STAF) and the measurement domain (M-STAF). This paper focused on the latter. M-STAF is a field-based framework; its fields are instrument-agnostic and enable cross-pollination between very diverse teams. We have shown how M-STAF (with its matrix tool and the requirement template) encourages consistency across instruments, completeness in the coverage of the requirements, and traceability of the engineering design all the way to the customer requirements. The tools that M-STAF suggests operate best when they enable structured conversations in a flexible framework, rather than being imposed as rigid formulas that must be followed. Throughout the paper we provided an example of the framework can be implemented and used on the planned Europa Mission as a case study.

In the end, the measurement Science Traceability and Alignment Framework in particular, and STAF in general, was developed to ensure that these large space missions serve the science that they are designed to advance. When engineers have tools to better communicate with their counterpart scientists, and share a deeper understanding of the science connections both up to the customer requirements and down to the subsystem design, the system design can be better harnessed to ensure the highest possible science return for the mission. And that is a sentiment that both scientists and engineers can share.

ACKNOWLEDGEMENTS

The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

The authors especially thank the Europa-UVS team, its PI Kurt Retherford, and its Investigation Scientist Scott Edginton; and the PIMS team, its PI Joseph Weslake, and its Investigation Scientists Corey Cochrane, and Abigail Rymer for allowing us to use their instruments as an example for the M-STAF implementation.

The authors also thank Professor Janet Vertesi for the fascinating and illuminating discussions on trading zones; Brian Cooke and John Day (Europa Mission Project System Engineer and Deputy respectively) for the support given in discussing the M-STAF; Robert Pappalardo, David Senske & Haje Korth, and Brian Paczkowski (Europa Mission Project Scientist, Deputy Project Scientists, and Science Manager) for all the dedication and patience demonstrated in the many discussions regarding the mission science; Jan Ludwinski, Eric Ferguson, Ben K. Bradley and Kelli McCoy, for embracing the idea of M-STAF and showing how to use it for a more efficient tour analysis.

REFERENCES

[1] L. L. Jones-Wilson and S. Susca, "A Framework for Extending the Science Traceability Matrix and its Application to the Planned Europa Mission," in *2017 IEEE Aerospace Conference* , Big Sky, MT, March 4, 2017.

[2] ISO/IEC/IEEE, "International Standard - Systems and software engineering -- Life cycle processes -- Requirements engineering," *ISO/IEC/IEEE 29148-2011,* pp. 1-94, 01 12 2011.

[3] NASA, "NASA Systems Engineering Handbook," Washington, D.C., 2007.

[4] IOC/IEC/IEEE, "International Standard for Systems and Software Engineering -- Life Cycle Management -- Part 4: Systems Engineering Planning," *ISO/IEC/IEEE 24748-4,* pp. 1-73, 15 05 2016.

[5] D. Mitchell, L. Lanzerotti, C. Kim, M. Stokes, G. Ho, S. Cooper, A. Ukhorskiy, J. M. ·. S. Jaskulek, D. Haggerty, P. Brandt, M. Sitnov, K. Keika, J. Hayes, L. Brown, R. Gurnee, J. Hutcheson, K. Nelson, N. Paschalidis and E. Rossano, "Radiation Belt Storm Probes Ion Composition Experiment (RBSPICE)," *Space Science Review,* vol. 179, pp. 263 - 308, 2013.

[6] R. Dodge, M. A. Boyles and C. E. Rasbach, "Key and Driving Requirements for the Juno Payload Suite of Instrument," in *AIAA Space 2007 Conference & Exposition*, Long Beach, CA, 2007.

[7] J. R. Weiss, W. D. Smythe and W. Lu, "Science Traceability," in *IEEE Aerospace Conference Proceedings*, April 2005.

[8] J. Sengal, "Scientists and software engineers: A tale of two cultures," in *Proceedings of the 20th Annual Meeting of the Pschology of Programming Interest Group*, Lancaster, UK, 2008.

[9] P. Galison, "Trading zone: coordinating action and belief," in *The Science Studies Reader*, M. Biagioli, Ed., New York and London, Routledge, 1999, pp. 137 - 160.

[10] P. Collopy and P. Hollingsworth, "Value-Driven Design," in *9th AIAA Aviation Technology, Integration, and Operations Conference*, Hilton Head, South Carolina, 21-23 September 2009.

11] D. Selva and E. F. Crawley, "VASSAR: Value Assessment of System Architectures using Rules," in *IEEE Aerospace Conference*, Big Sky, MT, 2013.

BIOGRAPHY

Sara Susca *received her PhD in electrical engineering from UCSB in 2007. She spent three years at Honeywell Aeropspace developing new technology for GPS denied navigation. She has been with JPL since 2011 where she covered various roles including project manager for STABLE (a* *balloon-borne sub-arcsecond pointing demonstration). She is currently member of the Europa Mission payload team as the instrument engineer for EIS and PIMS instruments.*

***Laura Jones-Wilson** earned her B.S. in Aerospace Engineering at Virginia Tech in 2007 and went on to study at Cornell University, where she obtained her M.S. and Ph.D. in Aerospace Engineering specializing in Dynamics and Controls. She then joined JPL in 2012 as a guidance and control systems engineer, where she served as an attitude control lead and systems engineer for a number of small projects before joining the Europa Mission payload team. She is currently serving as the instrument engineer for Europa-UVS.*

***Bogdan V. Oaida** is a Systems Engineer at JPL. He received a B.S.E. in Aerospace Engineering in 2007 and a M.Eng in Space Engineering in 2008, both from The University of Michigan. He has been the OPALS Project Systems Engineer over the entire lifecycle of the mission, including extended operations. In 2014 he received JPL's Early Career Achievement Explorer Award for his work on OPALS. He is currently a Payload Systems Engineer for the Europa Mission*